Class 13: Independence, Covariance, & Correlation

Thursday, October 17, 2019 11:39 AM

- HW 5 due Oct 24, 1pm (Thurs)
- Midterm grades are up

Recap:

Two discrete rv's X and Y are <u>independent</u> if

$$p_{X,Y}(x,y) = p_X(x)\,p_Y(y) \quad \text{for all } x, y.$$

Joint PMF = (PMF of X)(PMF of Y)

or

$$\mathbb{P}(X=x, Y=y) = \mathbb{P}(X=x)\,\mathbb{P}(Y=y).$$

Properties: Suppose X and Y are independent.

① Then $p_{X|Y}(x|y) = p_X(x)$ for all x, y, with $p_Y(y) > 0$.

Recall $p_{X|Y}(x|y) = \mathbb{P}(X=x \mid Y=y) = \dfrac{p_{X,Y}(x,y)}{p_Y(y)}$.

② $$\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$$

Note: $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

NOT true unless X and Y are independent.

↳ Ex: $X \sim \text{Ber}(\tfrac{1}{2})$, so $X = \begin{cases} 1 & \text{with prob. } \tfrac{1}{2} \\ 0 & \text{" } \tfrac{1}{2} \end{cases}$.

Let $Y = X$. Then $\mathbb{E}[XY] = \mathbb{E}[X^2] \overset{X^2 = X \text{ always}}{=} \mathbb{E}[X] = \tfrac{1}{2}$.

But $\mathbb{E}[X] = \tfrac{1}{2}$, $\mathbb{E}[Y] = \mathbb{E}[X] = \tfrac{1}{2}$, so

$$\mathbb{E}[XY] = \tfrac{1}{2} \neq \tfrac{1}{4} = \mathbb{E}[X]\,\mathbb{E}[Y].$$

③ For functions $f: \mathbb{R} \to \mathbb{R}$, $g: \mathbb{R} \to \mathbb{R}$,

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\,\mathbb{E}[g(Y)].$$

Proof:

$$\mathbb{E}[f(X)g(Y)] = \sum_x \sum_y f(x)g(y)\,p_{X,Y}(x,y)$$

$$= \sum \sum f(x)g(y)\,p_X(x)\,p_Y(y) \qquad \text{by independence}$$

$$= \sum_x \sum_y f(x) g(y)\, p_X(x) p_Y(y) \qquad \text{by independence}$$

$$= \left(\sum_x f(x) p_X(x)\right)\left(\sum_y g(y) p_Y(y)\right)$$

$$= \mathbb{E}[f(X)]\,\mathbb{E}[g(Y)]$$

Note: Independence does not help us compute $\mathbb{E}[h(X,Y)]$ unless $h(x,y) = f(x)g(y)$.

④ (Recall X and Y are independent.)

$$Var(X+Y) = Var(X) + Var(Y).$$

This needs X and Y to be independent.
(or at least uncorrelated, see later...)

Note: $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ without any independence assumption.

Proof:

Let's assume $\mathbb{E}[X] = \mathbb{E}[Y] = 0$. This is ok because, e.g.,

$$Var(X) = Var(\underbrace{X - \mathbb{E}[X]}_{\tilde{X}}) \quad \tilde{X} \to \mathbb{E}[\tilde{X}] = 0$$

(shifting X by a constant does not change variance)

Then

$$Var(X+Y) = \mathbb{E}\left[\left((X+Y) - \underbrace{\mathbb{E}[X+Y]}_{=0 \text{ since } \mathbb{E}[X]=\mathbb{E}[Y]=0}\right)^2\right]$$

$$= \mathbb{E}[(X+Y)^2]$$

$$= \mathbb{E}[X^2 + 2XY + Y^2]$$

$$= \mathbb{E}[X^2] + 2\underbrace{\mathbb{E}[XY]}_{=\mathbb{E}[X]\mathbb{E}[Y] \text{ by independence}} + \mathbb{E}[Y^2]$$

$$= \mathbb{E}[X^2] + 2\underbrace{\mathbb{E}[X]}_{=0}\underbrace{\mathbb{E}[Y]}_{=0 \text{ by assumption}} + \mathbb{E}[Y^2]$$

$$= \mathbb{E}[X^2] + \mathbb{E}[Y^2]$$

$$= \mathbb{E}[(X - \mathbb{E}[X])^2] + \mathbb{E}[(Y - \mathbb{E}[Y])^2]$$

$$= Var(X) + Var(Y)$$

Note: $Var(X) = \mathbb{E}[X^2]$ if $\mathbb{E}[X] = 0$.

$$= Var(X) + Var(Y)$$

$\underline{3 \text{ or more random variables}}$: same facts hold pretty much.

Let $X_1, X_2, \ldots, X_n$ be discrete rv's.

We say they are <u>(jointly) independent</u> if

$$\mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$
$$= \mathbb{P}(X_1 = x_1)\mathbb{P}(X_2 = x_2) \cdots \mathbb{P}(X_n = x_n),$$

for all $x_1, x_2, \ldots, x_n \in \mathbb{R}$.

<u>Properties</u>:

(1) $\mathbb{E}[f_1(X_1) f_2(X_2) \cdots f_n(X_n)] = \mathbb{E}[f_1(X_1)]\,\mathbb{E}[f_2(X_2)] \cdots \mathbb{E}[f_n(X_n)]$

(2) $Var(X_1 + X_2 + \ldots + X_n) = Var(X_1) + Var(X_2) + \ldots + Var(X_n)$.

<u>Ex 1</u>: Binomial Distribution. Let $0 < p < 1$, $n \in \mathbb{N}$.

Let $X_1, \ldots, X_n$ be independent $Ber(p)$.

Let $X = \sum_{i=1}^{n} X_i$. Then $X \sim Bin(n,p)$.

This gives us an easy way to calculate...

linearity of expectation

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} p = np$$

independence

and

$$Var(X) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$$
$$= \sum_{i=1}^{n} p(1-p) = np(1-p).$$

Recall: If $X_i \sim Ber(p)$ then $\mathbb{E}[X_i] = p$ and $Var(X_i) = p(1-p)$.

<u>Ex 2</u>: Let $X \sim Poisson(\lambda)$ and $Y \sim Poisson(\mu)$, $\lambda, \mu > 0$.

Then $p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}$ and $p_Y(k) = e^{-\mu}\frac{\mu^k}{k!}$.

Then $p_X(k) = e^{-\lambda}\frac{\lambda}{k!}$ and $p_Y(k) = e^{-\mu}\frac{\mu^k}{k!}$.

Assume X and Y are independent.

Let $Z = X+Y$.

Goal: Show $Z \sim \text{Poisson}(\lambda+\mu)$. (Transformation of joint distributions.)

Proof: The PMF of Z is

Note: $\{X+Y=n\} = \underbrace{\bigcup_{k=0}^{n}\{X=k, Y=n-k\}}_{\text{disjoint}}$

$$\begin{aligned}
\mathbb{P}(Z=n) &= \mathbb{P}(X+Y=n) \\
&= \sum_{k=0}^{n} \mathbb{P}(X=k, Y=n-k) \\
&= \sum_{k=0}^{n} \mathbb{P}(X=k)\,\mathbb{P}(Y=n-k) && \text{by independence} \\
&= \sum_{k=0}^{n} e^{-\lambda}\frac{\lambda^k}{k!}\, e^{-\mu}\frac{\mu^{n-k}}{(n-k)!} \\
&= e^{-(\lambda+\mu)} \sum_{k=0}^{n} \frac{1}{k!(n-k)!}\lambda^k \mu^{n-k} \\
&= e^{-(\lambda+\mu)} \frac{1}{n!} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\lambda^k \mu^{n-k} \\
&= e^{-(\lambda+\mu)} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k}\lambda^k \mu^{n-k} \\
&= e^{-(\lambda+\mu)} \frac{1}{n!} (\lambda+\mu)^n && \text{by binomial theorem} \\
&= \mathbb{P}(\text{Poiss}(\lambda+\mu) = n).
\end{aligned}$$

Covariance & Correlation

Def: If X and Y are any rv's, the covariance is defined by

$$\text{Cov}(X,Y) = \mathbb{E}\big[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\big].$$

Intuition: Positive if X and Y tend to "deviate" from their means in the same direction.

Facts:

① $Cov(X,X) = Var(X)$

② $Cov(aX+b, Y) = a\,Cov(X,Y)$, $a,b \in \mathbb{R}$ (Hw 5)

③ $Cov(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

Proof:

$$\begin{aligned} Cov(X,Y) &= \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] \\ &= \mathbb{E}[XY - Y\mathbb{E}[X] - X\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y]] \\ &= \mathbb{E}[XY] - \underline{\mathbb{E}[Y]\mathbb{E}[X]} - \mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y] \end{aligned}$$

$\mathbb{E}[Y\mathbb{E}[X]] = \mathbb{E}[Y]\mathbb{E}[X]$ since $\mathbb{E}[X]$ is constant

$$= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

④ If X and Y are independent, then $Cov(X,Y) = 0$.
Follows from ③ and $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

Note: We say X and Y are uncorrelated if $Cov(X,Y) = 0$.

Thus independence $\Rightarrow$ uncorrelated,

but uncorrelated does not imply independence! (Ex on Hw 5)

(If $f(X)$ and $g(Y)$ are uncorrelated for every f and g, then X and Y are independent.)

Note: Covariance has units!

But correlation does not. It is unitless.

Def: The correlation of X and Y is

$$\frac{Cov(X,Y)}{Sd(X)\,Sd(Y)} = \frac{Cov(X,Y)}{\sqrt{Var(X)\,Var(Y)}}.$$

Sometimes written $\rho(X,Y)$ or $Corr(X,Y)$. (Not as standardized.)

Facts:

(1) $-1 \le \rho(X,Y) \le 1$ always! $\quad |Cov(X,Y)| \le Sd(X)Sd(Y)$

Cauchy-Schwarz inequality

If $\rho = 1$ we say X and Y are perfectly correlated.

$\rho = -1$ we say X and Y are perfectly negatively correlated.

If $\rho > 0$, positively corr.

$\rho < 0$, negatively corr.

(2) $$\rho(aX+b, Y) = \begin{cases} \rho(X,Y) & \text{if } a > 0 \\ -\rho(X,Y) & \text{if } a < 0 \end{cases}$$

Ex 1: Flip 2 coins. $Y = \#$ heads, $X = \begin{cases} 1 & \text{if } 1^{st} \text{ flip heads} \\ 0 & \text{else.} \end{cases}$

Joint PMF:

$Y \sim Bin(2, \frac{1}{2})$ $\quad X \sim Ber(\frac{1}{2})$

X

0 1

$E[X] = \frac{1}{2}$, $\quad E[Y] = 1$

$Var(X) = \frac{1}{2}(1-\frac{1}{2}) = \frac{1}{4}$ $\quad Var(Y) = 2 \cdot \frac{1}{4} = \frac{1}{2}$

Y \ X	0	1
0	1/4	0
1	1/4	1/4
2	0	1/4

$\mathbb{E}[X] = \frac{1}{2}$, $\quad \mathbb{E}[Y] = 1$

$\mathrm{Var}(X) = \frac{1}{2}(1-\frac{1}{2}) = \frac{1}{4}$, $\quad \mathrm{Var}(Y) = 2 \cdot \frac{1}{4} = \frac{1}{2}$,

$\mathbb{E}[XY] = 1 \cdot 1 \cdot \frac{1}{4} + 1 \cdot 2 \cdot \frac{1}{4} = \frac{3}{4}$,

$\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{3}{4} - \frac{1}{2} \cdot 1 = \frac{1}{4}$,

$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} = \frac{1/4}{\sqrt{\frac{1}{4} \cdot \frac{1}{2}}} = \frac{1}{\sqrt{2}} > 0$ as expected!